Holistic Resource Management and Air Interface Abstraction Models

P. Karimi[1], F. Sanchez Moya[2], Ö. Bulakci[3], D. M. Gutierrez-Estevez[4], M. Ericson[5], A. Prasad[6], G. Fodor[7]

[1] WINLAB, Rutgers University, New Jersey, United States; [2]Nokia Bell Labs, Wroclaw, Poland; [3]Huawei Technologies GRC, Munich, Germany; [4]Samsung Electronics R&D Institute UK, Staines-upon-Thames, United Kingdom; [5]Ericsson Research, Luleå, Sweden; [6]Nokia Bell Labs, Espoo, Finland; [7]Ericsson Research, Stockholm, Sweden

***Abstract*—Resource Management (RM) in fifth generation (5G) radio access network has been the subject of extensive study in recent years. An agile RM framework should encompass greatly-improved services from previous mobile network generations along with advanced new services meeting the challenging requirements, like massive data rate and ultra-low latency. The novel aspects of 5G system lies within co-existence of diverse legacy or new wireless technologies and air interfaces (AI) consuming shared resources of frequency, time, etc. This will lead to new use-cases, which should holistically be addressed by a comprehensive RM framework. In this paper we explore a number of these use cases, namely flexible services-based time scheduling, integration and abstraction of multiple AI variants (AIV) in a systematic way. RM for device-to-device communication and RM for inter-network collaboration have been shown as examples exploiting this model abstraction.**

***Index Terms*—5G, flexible air interface, new radio, resource management, METIS-II**

I. Introduction

THE agile resource management (RM) framework which will be used in fifth generation (5G) systems should holistically consider the novel and differentiating aspects of 5G systems with respect to previous generations of mobile communication standards, specifically in terms of diverse and challenging services and use cases, existence of multiple air interface variants (AIVs), dynamic topologies, and novel communication modes (e.g., device-to-device communications, D2D). On this basis, the agile RM framework provides holistic RM solutions and air interface (AI) abstraction models that consider and exploit the novel aspects of 5G systems, such as, very diverse service requirements, existence of multiple AIVs in the overall 5G AI, dynamic topologies, and novel communication modes. Within the context of agile RM, the notion of a resource is extended beyond conventional radio RM (RRM) to attain the optimum mapping of 5G services to any available resources when and where needed within this extended realm of resources. In addition to the licensed radio frequency bands the extended realm of resources includes the unlicensed bands, whose usage shall be adaptive and be coupled with the changing radio topology, as well as hardware and software resources.

In this paper our focus is to provide holistic solutions to adhere by the diverse requirements of 5G use cases and services. To this end, it comprises the following main concepts: Flexible multi-service scheduling captured Section II, abstraction models for 5G AIVs captured in Section III, RM for inter-network collaboration and RM for 5G D2D captured in Section IV and Section V, respectively. In each section, each of these concepts are elaborated and detailed analysis of how the RM framework addresses the 5G use cases is illustrated. The last section concludes the paper.

II. Flexible Multi-Service Scheduling

It is well known from the existing literature that there are fundamental tradeoffs between scheduling users to maximize their spectral efficiency, coverage, latency, or reliability [1]. A possible solution to deal with these tradeoffs is to support scheduling with different transmission time intervals (TTI) sizes per user and per scheduling instance [2]. This allows to simultaneously accommodate very different service requirements, scheduling each user with a TTI duration according to its corresponding optimization target.

As a first step in the direction of evaluating the potential and performance of flexible multi-service scheduling with variable TTI size, an initial set of system-level simulation results is provided. It compares the influence on performance of several (fixed per simulation) TTI size configurations, which allows to gain insight into the most suitable TTI duration that should be dynamically chosen per user, depending on its service requirements, radio channel quality and system load conditions.

The evaluation is performed in a 3GPP Urban Macro scenario, with 3 sectors per base station, 500 m inter-site distance and 21 cells in the system [3]. In-resource control channel (CCH) scheduling grants with link adaptation are assumed, which allows to model different degrees of CCH overhead (i.e. aggregation levels or number of resource elements) depending on the user radio conditions [2][4]. The traffic model consists of a mix of MBB and low-latency traffic, with the former being modeled as a single user full buffer download and the latter, being higher priority, follows a Poisson arrival process with 1 kB payload and varying total cell offered load [4].

The analysis here focuses on the latency performance (i.e. MAC layer one-way user-plane latency) at different percentiles, under different load conditions and TTI sizes. The throughput performance can be found in [4]. In Fig. 1, it can be observed that at low system loads, using a short TTI (e.g. 0.25 ms) is an attractive solution to achieve low latency communications due to the low transmission delay required to serve the payloads. However, as the load increases, longer TTI configurations with lower relative CCH overhead (due to a higher number of symbols available for data, and therefore higher spectral efficiency) provide better performance as these can better cope with the non-negligible queuing delay.

The benefits of a long TTI become even more evident when looking at the tail of the latency distribution. Even at 4 Mbps load, a 0.5 ms TTI size offers better latency performance than the 0.25 ms TTI if the percentile of interest is above the 99%. The 1 ms TTI configuration is found to be beneficial from a latency point of view for high loads and above the 99.9 percentile. The main reason for this behavior is the queuing delay. As the offered load increases, the queuing delay becomes the most dominant component of the total latency, especially for users experiencing very low SINR. Therefore, it is beneficial to increase the spectral efficiency of the transmissions (by using a longer TTI) in order to reduce the experienced delay in the queue. The observed trends are relevant for uMTC use cases, which require latency guarantees of a few milliseconds with reliability levels up to 99.999%.

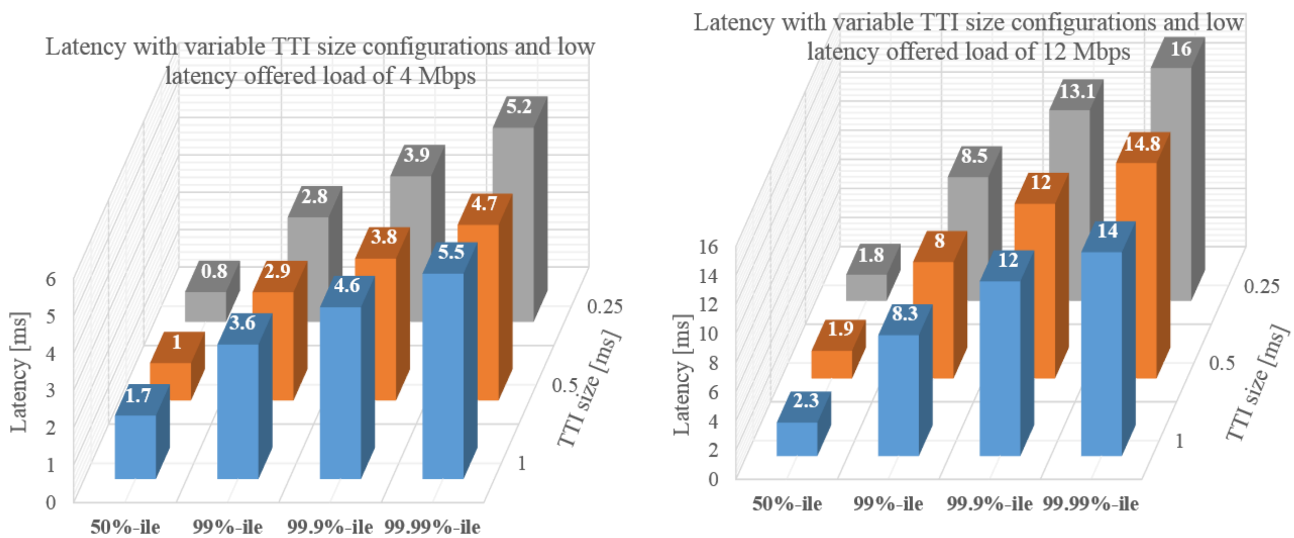

Figure 1. Latency (i.e. MAC layer one-way user plane latency) with variable TTI size configurations and low-latency offered loads.

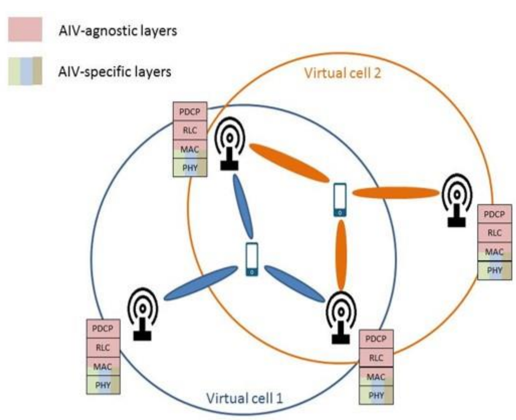

Figure 2. Virtual cell concept and AIV-agnostic versus AIV-specific layers.

These results, together with the findings in [13] (focused on TCP performance with variable TTI (Transmission Time Interval) size configuration), show the manifold benefits of having the flexibility to configure a flexible TTI size per user and per scheduling instance. The TTI size can also be selected according to the individual user's service requirements (e.g. besides uMTC and xMBB use cases, mMTC with narrow bandwidth operation would for instance benefit from longer TTI configurations).

III. Abstraction Models for 5G AIVs

The existence of multiple 5G AIVs requires the study of different integration options, determining what degree of AIV-specific versus AIV-agnostic RM functionalities is needed and at which level in the protocol stack. Ongoing work focuses on developing an abstraction framework for air interfaces where AIV-specific and AIV-agnostic features (e.g., frame structure, waveform, frequency band, etc.) and functionalities (e.g., RRM, interference management, scheduling, etc.) are identified and exploited in cooperative base station techniques. A user-centric virtual cell is proposed as the main architectural abstraction to achieve edgeless user experience in heterogeneous and dynamic topology scenarios, see Fig. 2. In contrast to static configurations with predefined central controllers, a user-centric virtual cell achieves this by utilizing a group of cooperating nodes wherein a user is served by one or more dynamically assigned nodes, and the virtual cell is continuously reformed trying to keep the user at the center of the cell. A primary goal of these virtual cells is to provide a uniform quality of experience to users anywhere in the system, by “eliminating” the edge, i.e., to provide uniform SINR, eliminate handover, and provide a sustained TCP throughput for a uniform service experience regardless of the user location.

TABLE 2 CURRENT D2D STATUS IN 3GPP VS. WANTED D2D FUNCTIONALITY IN 5G

Use case	**Status in 3GPP**	**5G D2D**
Unicast communication in cellular coverage (in licensed spectrum)	Not currently in 3GPP	Unicast D2D in cellular spectrum should be supported. Cellular layer must be protected
Group-cast/broadcast in cellular coverage (in licensed spectrum)	Available in 3GPP	Performance enhancements
D2D based relaying	Available in 3GPP (R13 on-going)	Performance enhancements
Cooperative devices in 5G coverage	Unavailable in 3GPP	Academic and industrial research available
D2D/Adhoc network outside 5G coverage	Available in 3GPP (for group-/broad-cast, no UE-UE relay)	All the above when no network assistance is available

IV. RM FOR 5G D2D

In LTE, D2D has been added as a feature on top of an already mature system, whereas 5G systems present the opportunity to natively support and capitalize on the advantages of D2D and other novel communication variants, such as self-backhauling. Therefore, a 5G holistic RM framework shall consider them. A basic support for broadcast based D2D communications was first added in LTE Release 12 [6]. The main functionalities were developed for the public safety (PS) use case, including intra- and inter-cell (in-coverage), outside network coverage and partial network coverage scenarios. For non-public safety use cases only discovery within network coverage was supported. For Release-13 and Release-14 the scope of D2D communications is extended both for PS and commercial use cases, including support for vehicle-to-anything (V2X) communications [7]. Still, some features with potential for 5G services and verticals are missing. Table 2 lists the current status of D2D in 3GPP and some of the desired D2D enhancements for 5G.

From a RRM perspective, network-controlled D2D shall allow to efficiently activate direct mode of communication among devices when needed (e.g. based on service requirements or system/user metrics) and effectively utilize all available resources to multiplex both D2D and cellular users.

V. RM FOR INTER-NETWORK COLLABORATION

The Software Defined Wireless Access Network (SDWN) approach has recently been studied and identified as a promising way to allow innovation and quicker evolution of the architecture [8][9]. While the corresponding transition to a software defined approach has now somewhat solidified in the networking world reaching wide spread commercialization, the same concepts are still evolving for wireless access networks. First and foremost, while this approach cleanly separates functionalities across elements and leads to easier-to-write, globally-optimum algorithms, it comes at the cost of scalability and latency. In particular, two fundamental issues arise which did not appear in the SDN reference world: a) First, SDWN requires a neighborhood view of the physical resources in order to take effective choices for common control parameters such as channel, power, rate, etc. b) Second, due to different time scale requirements generated by the various operations and functionalities that only apply to the wireless world (e.g. scheduling, MAC access mechanisms), a strictly centralized approach can quickly become unfeasible, leading toward the need of a split of the management operations in centralized and local (e.g. at the base station) ones.

a) Core Components

In order to approach the problem, two core components create the core of the collaboration framework.

Open Base Station: It is believed that the future 5G architecture will benefit from the concept of “open base stations” and “open access points” which interface to a common (“software defined wireless network”) CP that can be used to support a wide range of RRM and access network functionalities. The proposed open/software-defined wireless network approach has many advantages for 5G including the ability to support multiple radio air interfaces within the same framework while enabling a clean separation between PHY/MAC and networking or management functionality. Based on an ongoing research effort [10] aimed at developing a software-defined Control Plane framework for wireless networks, similar in spirit to SDN/OpenFlow standards developed for wired networks over the past few years, Open Base Station uses REST-based application programming interface (API) which is available through UP and facilitates common CP support across different Air Interfaces (i.e. WiFi, WiMAX, LTE and future 5G AIs).

SDN APIs and RadioMap: Towards the development of the Open Base Station concept, new APIs are defined that allow control of the most common parameters such as frequency, power and a concept of *RadioMap* information retrievable through such API. The *RadioMap* is in particular a powerful abstraction that allows for the collection of the previously discussed neighborhood view, necessary towards fully deploying coordination techniques in inter-network and inter-technology scenarios.

Further analysis will be dedicated towards understanding how to divide the aforementioned APIs across three core groups

based on the time scales of such functionalities:
A set of direct APIs (DC APIs) used to define explicit rules/actions to be performed by network elements (e.g.: UP setup, wireless access control, power control management at BS or UE).

A set of policy based APIs to be applied for locally decided operations (e.g.: scheduling). Statistics and events collection similar in spirit to counters in SDN. Fig. 3 shows such division in the architecture.

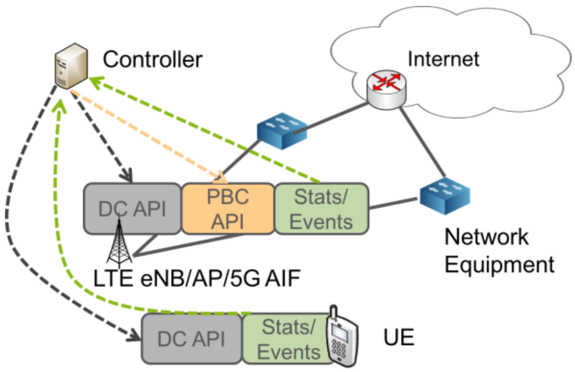

Figure 3. APIs for variety of air interfaces/common functions towards SDNization of wireless access networks.

b) Prototype-based Evaluation

Evaluation is being carried out using software prototypes deployed on the ORBIT wireless testbed [11] which enables realistic evaluation using software-defined radios and spectrum measurement.

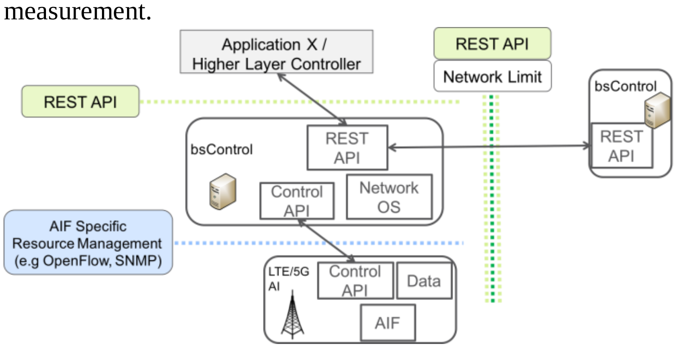

Figure 4. Prototype architecture diagram.

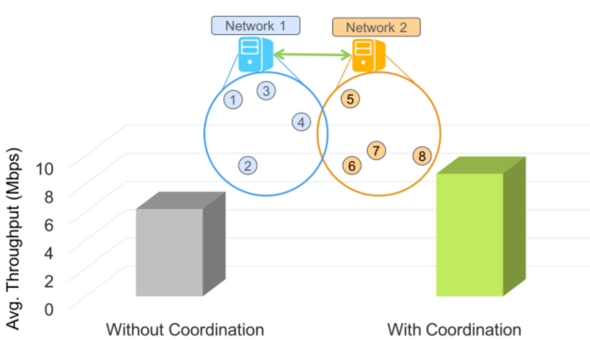

Figure 5. RadioMap Coordination Mechanisms Prototype Based Results.

The prototype consists of multiple software components extended to support external control APIs including the Open *AirInterface* – an open-source implementation of a LTE BS and UE – and *hostapd* – an open-source WiFi AP. Additionally, as shown in the block diagram depicted in Fig. 4, a custom controller, called *BSControl*, provides southbound control interfaces using hardware specific protocols (e.g. OpenFlow, SNMP, etc.) and that exposes a northbound REST API to control all exposable parameters of such hardware.

In order to evaluate the prototype, a baseline experiment focused only on Wi-Fi networks has been performed. The experiment consists of 8 Wi-Fi APs, each of which co-located with one Wi-Fi client.

In the first round of the experiment, labeled as "without coordination", each controller runs a graph coloring algorithm to assign channels to each AP-client pair using the algorithm in [12]. In the second round of the experiment, labeled as "with coordination", the two controllers share *RadioMap* information, here consisting of the location of the AP-client pair and the transmission power, and compute the channel allocation based on global information.

Fig. 5 shows the average throughput results for both scenarios. For this baseline scenario, a gain of more than 30% in average throughput is achieved when information is exposed compared to strictly applying the algorithm within each network.

Since the framework provides technology-agnostic abstractions, this baseline scenario can easily be extended to be used for multiple wireless networks, co-existing within the same region. An example of extending the channel allocation scenario for Wi-Fi and LTE coexistence is shown in Fig. 6. Here RadioMap can be extended to include necessary and aggregated frequency and time domain information, like power spectral density mask, start-stop frequencies, corresponding transmission power level, etc.

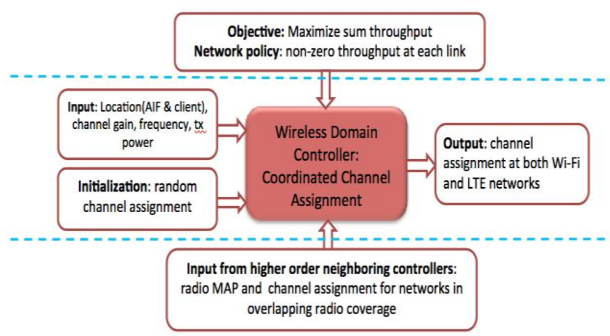

Figure 6. Channel allocation for wireless network using inter-network cooperation

VI. Conclusion

This paper presented some of the important use cases, which will holistically be addressed and solved within the agile RM framework in 5G systems. As a major step towards holistic solutions for novel services in 5G, we proposed abstraction models for various AIVs which will provide a technology-agnostic framework to mitigate the challenges of co-existence amongst the AIVs. Moreover, D2D communication and multi-service time scheduling have been discussed. All of these concepts were evaluated and result of analysis has been shown in each section.

ACKNOWLEDGMENT

This work has been performed in the framework of the H2020 project METIS-II co-funded by the EU. The views expressed are those of the authors and do not necessarily represent the project. The consortium is not liable for any use that may be made of any of the information contained therein.

REFERENCES

[1] B. Soret, et al., “Fundamental Tradeoffs among Reliability, Latency and Throughput in Cellular Networks,” IEEE Proc. GLOBECOM, December 2014.

[2] K.I. Pedersen, G. Berardinelli, F. Frederiksen, A. Szufarska, ”A Flexible 5G Frame Structure Design for Frequency-Division Duplex Cases”, in IEEE Communications Magazine, pp. 53-59, March 2016.

[3] 3GPP TR 36.814, “Further advancements for E-UTRA physical layer aspects,” Release 9, v9.0.0 (2010-03).

[4] G. Pocovi, et al., “On the Impact of Multi User Traffic Dynamics on Low Latency Communications”, accepted in ISWCS 2016.

[5] K.I. Pedersen, M. Niparko, J. Steiner, J. Oszmianski, L. Mudolo, S. Khosravirad, “System Level Analysis of Dynamic User-centric Scheduling for a Flexible 5G Design”, accepted in Globecom 2016.

[6] 3GPP TR 22.803, “Feasibility study for Proximity Services (ProSe),” Release 12, v.12.2.0 (2013-06).

[7] 3GPP TR 22.885, “Study on LTE support for Vehicle to Everything (V2X) services,” Release 14, v14.0.0 (2015-12).

[8] Ali-Ahmad, H., Cicconetti, C., La Oliva, A., Mancuso, V., Reddy Sama, M., Seite, P., & Shanmugalingam, S., “An SDN-based network architecture for extremely dense wireless networks”, in IEEE SDN for Future Networks and Services (SDN4FNS) 2013 (pp. 1-7).

[9] Brief, ONF Solution. "OpenFlow™-Enabled Mobile and Wireless Networks." (2013).

[10] Gautam Bhanage, Ivan Seskar, Dipankar Raychaudhuri, *A virtualization architecture for mobile WiMAX networks*, ACM SIGMOBILE Mobile Computing and Communications Review, v.15 n.4, October 2011.

[11] Raychaudhuri, Dipankar, et al. "Overview of the ORBIT radio grid testbed for evaluation of next-generation wireless network protocols." IEEE Wireless Communications and Networking Conference (WCNC), 2005.

[12] Mishra, Arunesh, Suman Banerjee, and William Arbaugh. "Weighted coloring based channel assignment for WLANs." ACM SIGMOBILE Mobile Computing and Communications Review 9.3 (2005): 19-31.

[13] A. Prasad, F.S. Moya, M. Ericson, R. Fantini, Ö. Bulakci, “Enabling RAN Moderation and Traffic Steering in 5G Radio Access Networks,” 84th IEEE Vehicular Technology Conference, Montreal, Sept. 2016.